T1 MAPPING AND IMAGE RECONSTRUCTION: EXAMINATIONS FOR QUANTITATIVE MRI

by
Zhen Hu

A thesis submitted to The Johns Hopkins University in conformity
with the requirements for the degree of Master of Science in Engineering

Baltimore, Maryland
May, 2021

Abstract

Magnetic resonance imaging (MRI) is a major imaging modality widely used in medical practice. Compared to other medical imaging techniques, MRI has several advantages such as non-invasiveness, high tissue contrast, additional information, etc. While most other imaging techniques directly generate medical images, a special characteristic of MRI is that raw data are acquired in the k-space, an extension of the Fourier space, which represents spatial frequency information. Therefore, reconstruction algorithms are needed to convert collected raw data to human-readable images of the scanned region.

Quantitative MRI refers to MRI techniques that map meaningful properties of tissues. Amongst other tissue-specific parameters, we are particularly interested in those related to *relaxation*, i.e., the process of protons returning to the original state from an excited state. There are two relaxation mechanisms in MRI, namely *transverse relaxation* and *longitudinal relaxation*, concerned with the transverse and longitudinal components of net magnetization of protons, respectively.

This thesis consists of three parts. Firstly, we introduce hardware and fundamental principles of MRI. In the second part, we define three tissue-specific parameters related to relaxation—T1 (for longitudinal relaxation), T2 (for transverse relaxation) and T2* (for reduced transverse relaxation). In particular, we focus on T1 and present several commonly used *T1 mapping* methods for calculating T1 values, such as the dual-flip-angle (DFA) method. The last part of the thesis serves as a tutorial that introduces several basic steps that reconstruct an image from acquired raw data.

Thesis Readers

Dr. Qin Qin (Primary Advisor)
Associate Professor
Department of Radiology and Radiological Science
Johns Hopkins University

Dr. Jiadi Xu
Associate Professor
Department of Radiology and Radiological Science
Johns Hopkins University

Dr. Jun Hua
Associate Professor
Department of Radiology and Radiological Science
Johns Hopkins University

This thesis is dedicated to Yuchong Pan, without whom I can sleep 3 more hours every day.

Contents

List of Figures

Chapter 1

MRI Fundamentals

1.1 Biomedical Imaging Modalities

There are four major biomedical imaging modalities—radiographic imaging, nuclear medicine, ultrasound and magnetic resonance imaging (MRI). In this section, we briefly introduce principles and characteristics of these modalities. Radiographic imaging, including X-ray and computed tomography (CT), transmits X-rays through the body and detects the X-rays on the other side of the patient. Nuclear imaging, to which planar scintigraphy, single-photon emission computerized tomography (SPECT) and positron emission tomography (PET) belong, injects radiotracers in the bloodstream of the body and detects gamma rays emitted from the radiotracers within the body. The third category, ultrasound, sends ultrasonic waves towards the body and then receives the reflected ultrasonic waves from the body. Lastly, MRI causes spin systems to precess with the patient in a strong magnetic filed, transmits radio frequency energy to the body, and receives radio frequency energy induced in the body.

The electromagnetic spectrum of different biomedical imaging modalities is shown in Figure 1-1. Among all imaging modalities, ultrasound has the lowest frequency of about 10^4 Hz, slightly higher than the highest frequency that a person can hear. MRI uses radio frequency, which is about 100 MHz. Optical imaging uses infrared and visible light. With higher frequency levels, X-ray and gamma ray are typically

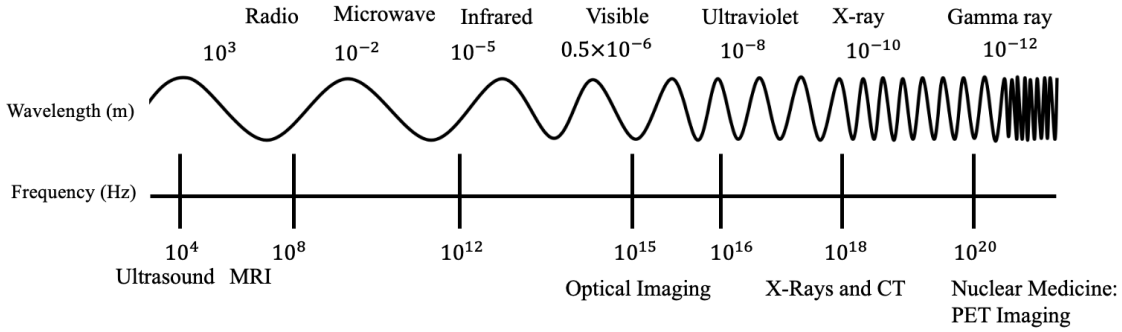

Figure 1-1. The electromagnetic spectrum for different biomedical imaging modalities.

used in CT and nuclear imaging, respectively.

Biomedical modalities can be categorized according to different criteria. X-rays, CT, SPECT and PET cause some level of radiation exposure to the patient, but they are still considered non-invasive. Ultrasound and MRI, on the other hand, are completely non-invasive. In terms of forms of imaging, CT, SPECT, PET, ultrasound and MRI belong to tomographic imaging, whereas X-rays and planar scintigraphy are projection imaging techniques. Moreover, SPECT, PET, diffusion-weighted imaging (DWI), perfusion-weighted imaging (PWI), and functional MRI provide information beyond anatomical information, such as physiological information, metabolic information, and functional information of the patient.

MRI has several advantages over other imaging modalities. As mentioned before, MRI is completely non-invasive and thus prevents radiation exposure of the patient. Secondly, it belongs to tomographic imaging, intrinsically providing additional information than projection imaging methods. Moreover, MRI provides relatively high spatial resolution and high soft tissue contrast. Another unique characteristic of MRI is the capability of scanning the subject along any direction. In this thesis, we review preliminary knowledge required for MRI techniques and then transition to two main areas of MRI—T1 mapping and image reconstruction.

1.2 MRI Hardware

In order to better understand how MRI works, we split the acronym "MRI" into three parts: magnetism, resonance and imaging. The magnetism part indicates the fact that the main magnet generates a strong magnetic field when a subject is inside the MRI scanner. The resonance part reflects the process of transmitting radio frequency (RF) energy to the subject by an RF coil, turning off the transmitter, and then receiving RF signals emitted by the subject using another RF coil (or sometimes the same RF coil). The imaging part means spatially modulating the magnetic field strength to distinguish signals from different locations using gradient coils.

There are three major hardware components of MRI, corresponding to the three aforementioned parts: a magnet, radio frequency transmitter and receiver (RF coils), and three set of gradient coils (along the x-, y- and z-directions, respectively). Specifically, the outer layer of an MRI scanner is the main magnet, with the three gradient coils placed inside the magenet, and with the RF coils being the inner core.

The magnet is typically superconducting, generating a magnetic field. The *Tesla* is typically used as the unit of the magnetic field strength in MRI, with 1 Tesla = 10,000 Gauss. The strength of the Earth's magnetic field ranges from 0.25 to 0.65 Gauss, and is about 0.5 Gauss in average. Hence the magnetic strength of a 1T MRI scanner is about 20,000 times that of the Earth. Particles inside the patient precess at a frequency proportional to the magnetic field strength due to the existence of the superconducting magnet. This frequency is called the *resonance frequency*.

The radio frequency transmitter and receiver, obviously, transmit and receive RF and induced signals, respectively. Protons that have the same precession frequency as the frequency in the RF transmitter receive energy and get excited. This phenomenon is called *magnetic resonance* and explained in detail in Section 1.3.

The three sets of gradient coils modulate the original magnetic filed spatically to

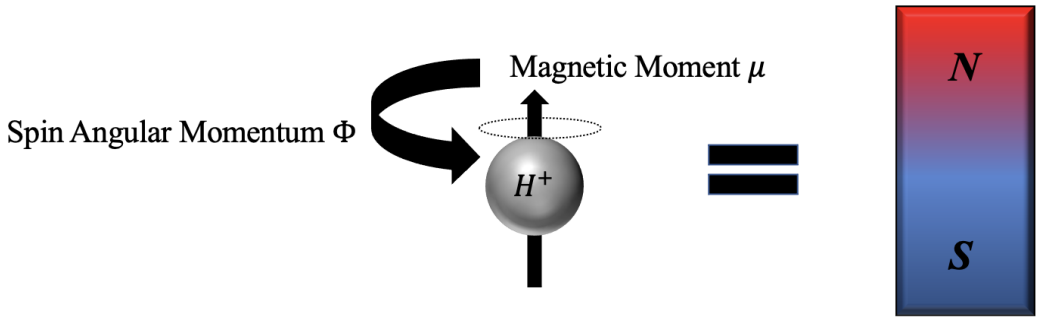

Figure 1-2. A proton behave like a tiny magnet.

provide spatial information for imaging. Without lost of generality, we assume that the magnetic field is always along the z-axis throughout this thesis. The gradient coils merely cause variations of the original field (along the z-direction) in different xy-positions but do not generate fields along the x- or y-direction. Therefore, the precession frequency ω of protons is a function of the spatial location. In contrast to other imaging modalities (e.g., CT, X-ray, SPECT, PET, Ultrasound, optical imaging), positional information of detectors (e.g., the RF coils in MRI) is not used to get spatial information in MRI except for parallel imaging.

1.3 Magnetic Resonance Phenomenon

1.3.1 Signal Source of MRI

A nucleus with an odd atomic number serves as a candidate for the signal source in MRI. In general, more than 99% of MRI scanners use protons (hydrogens, ^{1}H) as the signal source [1–3]. Because 70% of our body consists of water, a molecule of which has two hydrogen atoms, there are abundant signal sources in our body for MRI. There are also many other nuclei that can be used for MRI such as ^{13}C, ^{19}F, ^{23}Na, ^{17}O, ^{31}P [1, 2, 4].

The movement of a net charge within a closed loop is called *spin*, which generates an angular momentum Φ. The angular momentum generates a magnetic dipole moment

μ oriented orthogonal to the plane of the moving charge [5]. A proton in this situation behaves like a tiny magnet [2, 4] as shown in Figure 1-2. The spin angular momentum and the magnetic moment are related through a proportionality constant γ given by

$$\mu = \gamma \cdot \phi. \tag{1.1}$$

The constant γ is called the *gyromagnetic ratio*, which is 10^6 radians/sec/Tesla for a proton. Conventionally, often $\gamma/(2\pi)$ is used. Recall that Φ and μ are vector quantities. MR signals are generated by manipulations of this magnetic moment μ through "magnetic resonance".

1.3.2 Protons in Magnetic Field

Without an external magnetic field, all proton spins in our body are aligned randomly, as shown in Figure 1-3. However, when protons are placed in a strong magnetic field, two aspects of the magnetic moments can be observed:

1. The magnetic moments tend to be aligned parallel or anti-parallel to the external magnetic field (denoted by B_0), as shown in Figure 1-4.

2. The magnetic moments are induced to rotate around the axis of the magnetic field. This phenomenon is called *precession*. In other words, protons now spin and precess around the same axis, namely the external magnetic field direction.

Whether the magnetic moment of a proton is parallel or anti-parallel to B_0 depends on its energy. A proton whose magnetic moment is parallel to the field is in a low energy state whereas one with anti-parallel magnetic moment is in a high energy state, as shown in Figure 1-5. The number of protons parallel or anti-parallel to B_0 are approximately equal, so most of them cancel out each other in the net effect. However, the number of protons parallel to the field is slightly higher than the number of anti-parallel protons. This phenomenon is called *spin excess*, which is essential for generating MR signals.

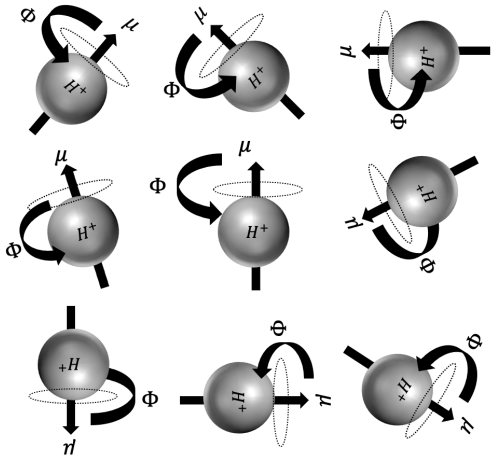

Figure 1-3. Protons are aligned randomly when there is no external magnetic field.

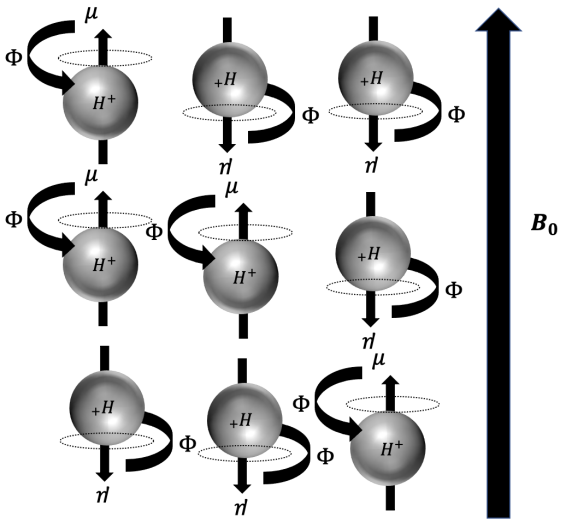

Figure 1-4. Magnetic moments tend to be aligned parallel or anti-parallel to the external magnetic field B_0.

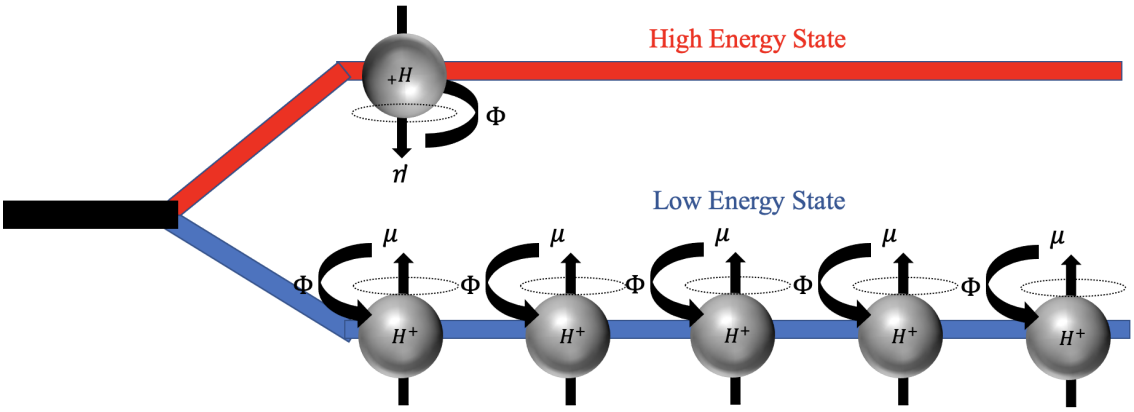

Figure 1-5. Protons in high energy states and in low energy states.

The precession of a proton is similar to that of a gyroscope. Without an external magnetic field, a proton is analogous to a vertically-oriented gyroscope spinning without wobbling. However, once the axis is tipped off, gravity causes the spinning gyroscope to wobble, similarly to the precession phenomenon when an external field exists. The wobbling, also known as *precession* has its own frequency, which is indepdendent of the spinning frequency.

The individual dipole moment of a proton is too small to detect. Instead, we consider the net effect of all dipoles. Each proton's magnetization can be regarded as a vector. Net magnetization of a precessing proton can be found using vector summation [1, 2]. Now we decompose a vector into a longitudinal component and a transverse component, only the latter of which can be used to generate MR signals. For the longitudinal component, the majority of parallel and anti-parallel protons cancel out each other, whereas few additional parallel protons called *spin excess* remain. As for the transverse component, the phases of protons are random, so the vector sum equals zero. Overall, net magnetization M is along the longitudinal direction (i.e., the z-direction), the same direction as the exteral field B_0.

Similar to the relation between the spin angular momentum and the magnetic moment in Equation (1.1), the precession angular speed ω_0 is proportional to the

exteral field B_0 with the proportional constant gyromagnetic radio γ, given by

$$\omega_0 = \gamma \cdot B_0, \tag{1.2}$$

Typically both sides of the equation are divided by 2π, transforming the angular speed to the corresponding regular frequency in Hertz, given by

$$f_0 = \frac{\gamma}{2\pi} \cdot B_0. \tag{1.3}$$

Equation (1.3) is called the *Lamor equation* and f_0 is called the *Larmor frequency*, which is usually assumed to be a constant for a given spin system. In 1T MRI, $f_0 = 42.58$ MHz [1–3, 6].

1.3.3 RF for Causing Magnetic Resonance

In this section, we introduce the magnetic resonance phenomenon from two different points of view—microscopic quantum mechanics and net magnetization. First, we explain the microscopic quantum mechanical interpretation. When the RF energy (often denoted as B_1; note that B_0 is the main magnetic field) with the same frequency as the precession frequency (the Larmor frerquency) applied to the transmitter RF coil, some protons absord energy and jump from a low energy state (parallel) to a high energy state (anti-parallel), as shown in Figure 1-6.

The other way to understand magnetic resonance is from the concept of net magnetization. Before we delve into the concept, we first introduce a new coordiante system, the *rotating reference frame*, for visualizaing magnetization dynamics. The net magetization itself precesses at the Larmor frequency. If we use the traditional laboratory reference frame, the precessing axis of the magnetic vector will constantly change, as shown in Figure 1-7. The rotating reference frame, however, precesses at the Larmor frequency so the net magetization is stationary. The rotating reference

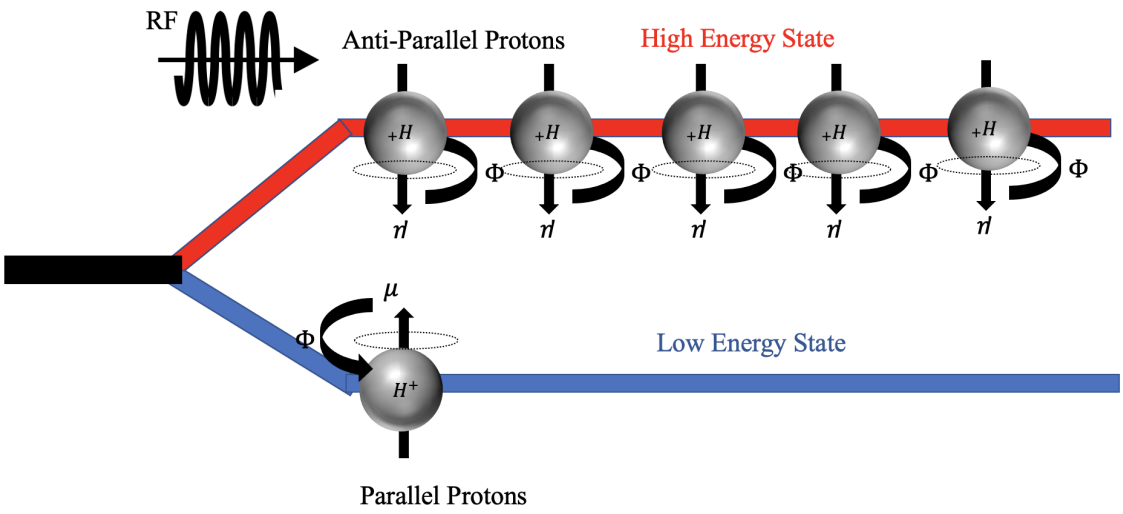

Figure 1-6. The proton states after applying the RF energy.

frame is expressed in Equation (1.4) and shown in Figure 1-8.

$$\begin{aligned} x' &= x \cdot \cos(\omega_0 t) - y \cdot \sin(\omega_0 t), \\ y' &= x \cdot \sin(\omega_0 t) + y \cdot \cos(\omega_0 t), \\ z' &= z. \end{aligned} \tag{1.4}$$

Now we are ready to introduce the magnetic resonance phenomenon. Figure 1-9 shows the entire process in the rotating reference frame. (If the laboratory frame were used instead, the trajectory would be in a spiral form as shown in Figure 1-10). The first subplot in Figure 1-9 illustrates spin excess. When we apply an RF pulse in the $x'y'$-transverse plane at the frequency of the precession frequency, the RF pulse decreases the overall net magnetization in the longitudinal direction M_z because these spins parallel to the field are changed to the anti-parallel state. The net transverse magnetization M_{xy}, which can be detected and measured by MRI scanners, starts to emerge (i.e., becomes in-phase), as shown in the second subplot in Figure 1-9. Specifically, when an RF pulse is applied along the x'-direction, the net magetization is pushed down toward the y'-axis accoring to the right-hand rule. If we apply an RF pulse continuously until the longitudinal component completely disappear, then the transverse component gets maximized. This RF pulse is called the 90° *RF pulse*, since

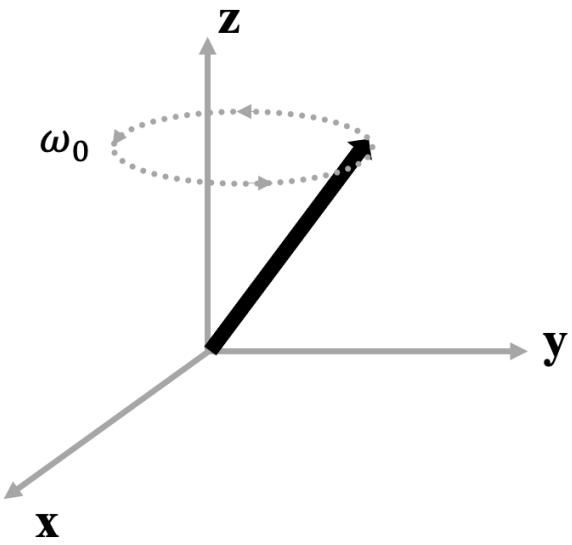

Figure 1-7. The laboratory reference frame, in which magnetic moment is rotating.

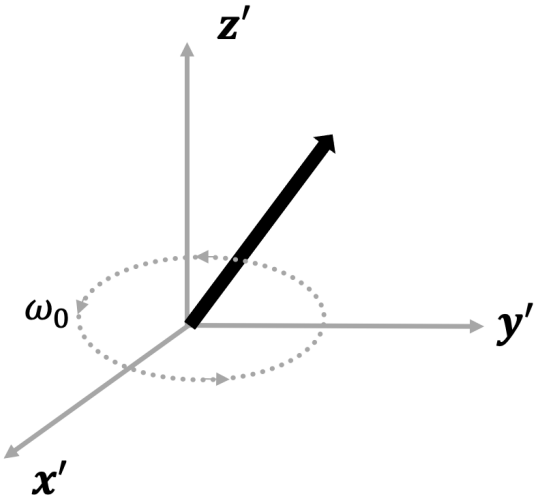

Figure 1-8. The rotation reference frame, which is itself rotating.

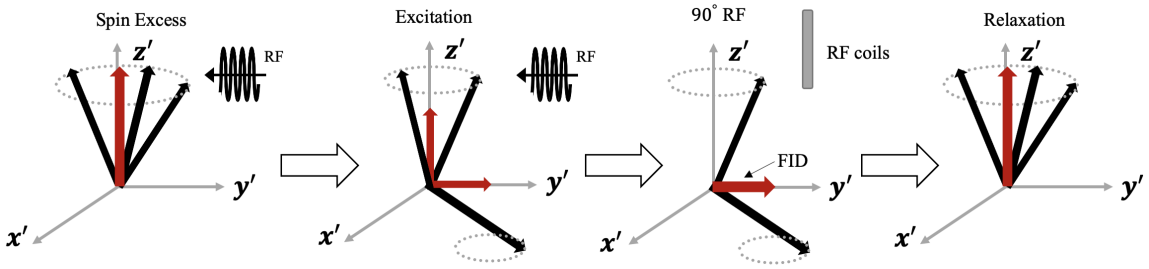

Figure 1-9. An illustration of the entire process of magnetic resonance.

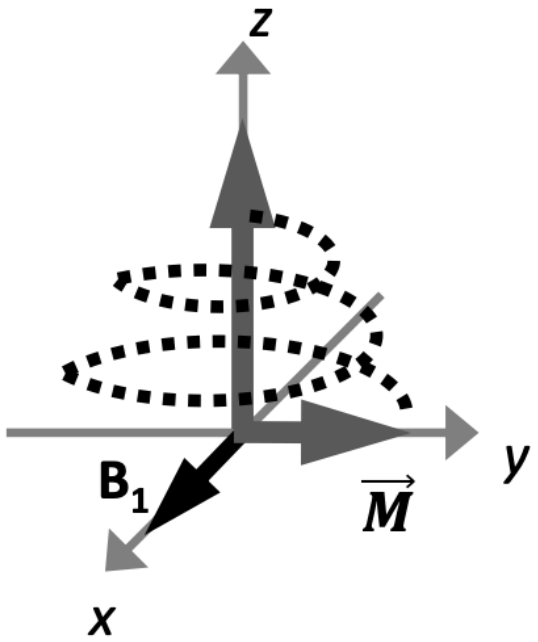

Figure 1-10. The trajectory of magnetic vectors in the laboratory reference frame.

the original longitudinal magnetization is flipped by 90° and hence completely toward the transverse plane. When we stop applying the RF pulse to the transmistter RF coil, we can detect an exponentially-decaying sinusoidal signal in the receiver RF coil, called *free induction decay (FID)*, as shown in the third subplot of Figure 1-9. Because M_{xy} is produced by multiple magnetic dipoles, differences in the microenvironment alter the speed at which these dipoles precess. Since some dipoles precess faster than others, they appear to spread apart in a process called *dephasing*. As these dipoles spread, they slowly begin to counteract with each other, reducing the net measurable M_{xy} [5]. The net magnetization eventually returns to the original state. This process is called *relaxation*, as shown in the final subplot of Figure 1-9.

If the duration of the 90° RF pulse is doubled, the net magnetization M is flipped to the $(-z)$-direction. This RF pulse is called the 180° *RF pulse*. If the RF energy B_1 were apply along the z-direction, it would merely change the precession speed yet not cause magnetic resonance. Therefore it is important to place RF coils so that

the RF fields generated by the RF coils are perpendicular to B_0. If the frequency of the RF pulse does not match the Larmor frequency, then no measurable signal in the transverse plane is produced.

The final flip angle, the angle between the z'-axis (the original net magetization location) and the eventual net magnetization location, and the phase of the net magetization M depend on both the amplitude and the duration of the RF pulse B_1 [2, 6], given in the following equation:

$$B_1(t) = \begin{cases} B_1^e(t)e^{-j(w_0 t-\theta)} & \text{in the laboratory frame,} \\ B_1^e(t)e^{j\theta} & \text{in the rotating frame,} \end{cases} \tag{1.5}$$

where B_1^e denotes the envelope of B_1.

If we apply the RF pulse along the x'-direction, the precession of the net megetization M around the x'-axis can be described by the Larmor equation $\omega_1 = \gamma B_1(t)$. In addition, the intergral of ω_1 over time provides the flip angle α given by

$$\alpha = \gamma \int_0^\tau B_1^e(t)dt,$$

where τ is the duration of the RF energy $B_1(t)$. The RF pulse is also known as the α *pulse*.

Chapter 2

T1 Mapping Simulation

2.1 Main Tissue Parameters

As mentioned in Section 1.3.3, applying an RF pulse (with $\alpha \neq \pi$) generates the transverse magnetization that can be detected as an exponentially-decaying sinusoidal signal (i.e., *free induction decay*) in an RF coil, as shown in Figure 2-1. Then the magnetization returns back to the original state, and this process is called *relaxation*. There are two relaxation mechanisms, *transverse relaxation* (also known as *T2 relaxation* or *spin-spin relaxation*) and *longitudinal relaxation* (also known as *T1 relaxation* or *spin-lattice relaxation*). In this section, we introduce three main tissue parameters: T1, T2, T2*.

Little information could be obtained if the response to RF pulses of protons were independent of its surrounding environment. Fortunately, the response of protons is indeed affected by its neighboring macromolecules and elements. Therefore, MR signals are produced by these multiple tissue-specific parameters [7, 8].

2.1.1 T1: Spin-Lattice Relaxation Time

The *T1 relaxation time*, or the *spin-lattice relaxation time*, measures how quickly the longitudinal component of the net magnetization M_z returns to its original state (or a percentage of the equilibrium) [1, 4, 6, 9–12], as shown in Figure 2-2. The energy

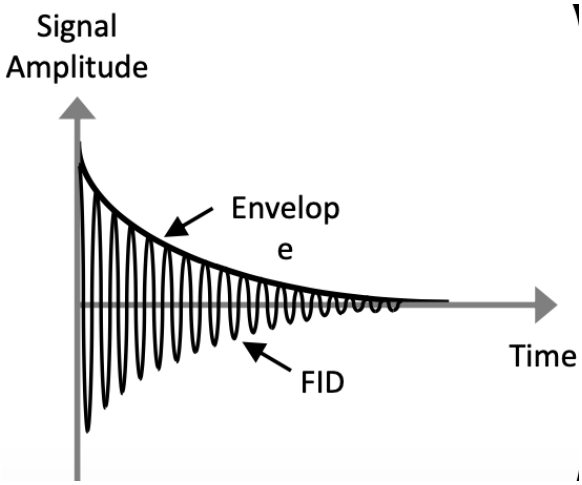

Figure 2-1. The free induction decay.

absorbed from the RF pulse during the excitation exchanges between protons and their surrounding environment in relaxation [8]. Returning from a high energy state to a low energy state emits energy to the surrounding neclei called the *lattice*, which serve as an energy reservoir.

2.1.2 T2: Spin-Spin Relaxation Time

The *T2 relaxation time*, also known as the *spin-spin relaxation time*, measures the time required for the transverse net magnetization M_{xy} to decay to a percentage of the initial magnitude strength, usually $1/e$ or 37%, due to the progressive dephasing of spinning dipoles after excitation. The so-called "spin-spin" interaction is one of several mechanisms that contribute to T2 relaxation. As we mentioned in T1 relaxation, when protons interact with their surrounding fluctuating magnetic fields, if such interaction affects one of the spins contributing to M_{xy}, both the transverse and the longitudinal components of its angular momentum will be randomly changed, so the progressive loss of phase relations with other spins will lead to the occurrence of T2 relaxation. Hence, T2 is partially dependent on T1 effects [4, 8, 10, 12]. In other words, any process causing T1 relaxation also results in T2 relaxation. In typical circumstances, the magnitude order of T2 is smaller than that of T1.

2.1.3 T2*: Reduced Transverse Relaxation Time

In reality, measured MR signals decay faster than the natural T2 decay because of magnetic field inhomogeneities [8, 13, 14]. The reduced T2, also known as T2*, is a derived tissue parameter that reflects this phenomenon. In addition to non-uniform external fields, there is one more possibility that results in field inhomogeneities—neighboring fields from adjacent tissues, commonly seen at interfaces of tissues [3, 4, 6, 13, 14], which also lead to the early dephasing of individual magnetic dipoles [15]. T2* is smaller than T2 as shown in Figure 2-3 and can be determined by the following equation

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'}, \tag{2.1}$$

where $T2'$ is the decay associated with magnetic field inhomogeneities.

$T2'$ relaxation is reversible by applying the 180° pulse, whereas T2 relaxation (i.e., the transverse relaxation) is irreversible. We explain the reason using an analogy. We first discuss the $T2'$ case. A group of students begins running from a starting line at the same time with possibly different speeds. The precession speeds are analogous to the students' speeds, related to the magnetic field strength according to the Larmor Equation (1.2). Due to magnetic field inhomogeneities, the precession speed is different for each proton. After a period of time, the students are at differnet locations, analogous to the fact that protons are dephased. If students are asked to return to the starting line using the same speeds as they come, they will arrive at the starting line at the same time. Analogously, if we apply a 180° RF pulse, the protons become in phase again, so $T2'$ relaxation procedure is reversible. Now let us consider the T2 scenario. If these students interact with each other randomly, they may arrive at slightly different positions compared to the $T2'$ situation. In this case, if we ask them to return, they may not arrive at the starting line at the same time due to random interactions. In other words, the interactions in the forward trip and

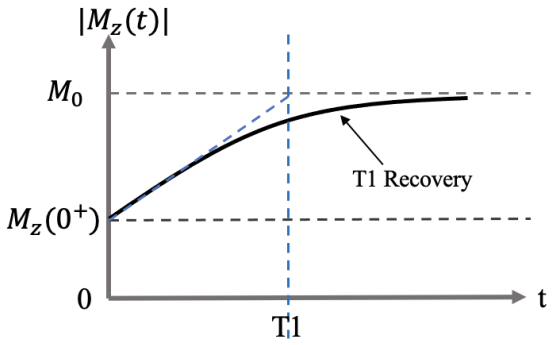

Figure 2-2. An illustration of T1 value and the T1 relaxation.

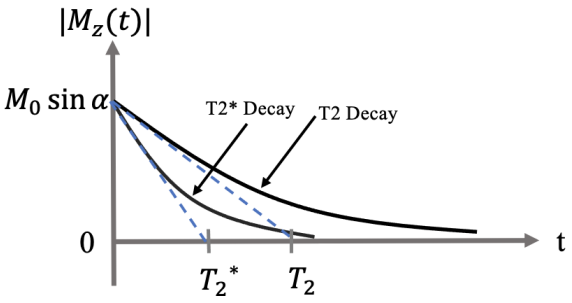

Figure 2-3. The comparison between the T2* decay and the T2 decay.

the backward trip may not be the same. The dephasing due to random interactions among the students is similar to T2 relaxation, which is irreversible.

The relaxation time constants T1 and T2 are tissue-specific and serve as sources of MR image contrasts, depending only on B_0 but not on scan parameters. T2* and $T2'$ are also tissue-specific but may depend on scan parameters. In subsequent sections, we focus on T1 relaxation and the classic T1 mapping methods.

2.2 Bloch Equation

In his 1946 paper *Nuclear Induction*, Bloch [10] showed a set of equations explaining the origin and the properties of nuclear magnetic resonance (NMR) signals. In this section, in order to better understand T1 relaxation, we derive the T1 decay formula from the basic Bloch equation; i.e., given $\frac{dM_z(t)}{dt} = \frac{M_0 - M_z(t)}{T_1}$, we derive the solution of $M_z(t)$. The M_0 is the equilibrium magnetization. By the general solution to a first

order non-homogeneous linear ordinary differential equation (derived in Appendix A), given

$$\frac{dy}{dt} = P(t)y + Q(t),$$

we have that

$$y = e^{\int P(t)dt}\left(\int Q(t)e^{-\int P(t)dt}dt + C\right). \tag{2.2}$$

In our case,

$$y = M_z(t), \qquad P(t) = -\frac{1}{T_1}, \qquad Q(t) = \frac{M_0}{T_1},$$

Therefore,

$$\begin{aligned} y = M_z(t) &= e^{\int -\frac{1}{T_1}dt}\left(\int \frac{M_0}{T_1}e^{-\int -\frac{1}{T_1}dt}dt + C\right), \\ &= e^{-\frac{t}{T_1}}\left(\int \frac{M_0}{T_1}e^{\frac{t}{T_1}}dt + C\right), \\ &= e^{-\frac{t}{T_1}}\left(M_0 e^{\frac{t}{T_1}} + C\right), \\ &= M_0 + Ce^{-\frac{t}{T_1}}. \end{aligned}$$

We substitute the initial condition $M_z(0)$ into this general solution to derive the constant C:

$$M_z(0) = M_0 + C.$$

Therefore,

$$C = M_z(0) - M_0. \tag{2.3}$$

The final solution is given by

$$M_z(t) = M_0 + (M_z(t) - M_0)e^{-\frac{t}{T_1}}. \tag{2.4}$$

Now we express Equation (2.4) in the matrix form $M(t) = AM(0) + b$.

$$\begin{bmatrix} M_x(t) \\ M_y(t) \\ M_z(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-\frac{t}{T_1}} \end{bmatrix} \begin{bmatrix} M_x(0) \\ M_y(0) \\ M_z(0) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ M_0(1 - e^{-\frac{t}{T_1}}) \end{bmatrix}. \tag{2.5}$$

Therefore,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-\frac{t}{T_1}} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ M_0(1 - e^{-\frac{t}{T_1}}) \end{bmatrix}. \tag{2.6}$$

Similarly, the T2 decay formula is given by

$$M_{xy}(t) = M_{xy}(0)e^{-\frac{t}{T_2}}, \qquad \text{or} \qquad M_{xy}(t) = M_{xy}(0)e^{-\frac{t}{T_2^*}}. \tag{2.7}$$

2.3 T1 Mapping Methods

Five T1 mapping methods are introduced in this section, including saturation recovery (SR), inversion recovery (IR), Look-Locker (LL), driven-equilibrium single-pulse observation of T1 (DESPOT1) and dual-flip-angle (DFA).

For saturation recovery, a 90° pulse is given initially, which gives $M_z(0) = 0$. Therefore, Equation (2.4) becomes

$$M_z(t) = M_0(1 - e^{-\frac{t}{T_1}}), \tag{2.8}$$

where M_0 is the equilibrium magnetization (equivalently, initial magnetization).

For inversion recovery, a 180° pulse is given initially, which gives $M_z(0) = -M_0$. Therefore, Equation (2.4) becomes

$$M_z(t) = M_0(1 - e^{-\frac{t}{T_1}}) - M_0 e^{-\frac{t}{T_1}} = M_0(1 - 2e^{-\frac{t}{T_1}}). \tag{2.9}$$

For Look-Locker, the longitudinal magnetization is described by

$$M_z(t) = M_{ss} - (M_{ss} + M_0)e^{-\frac{t}{T_1^*}}, \tag{2.10}$$

where M_{ss} is the longitudinal magnetization in the steady state, and T_1^* is the longitudinal magnetization relaxation time. Assuming that the transverse magnetization decay is ignorable (i.e., $T_2 \ll \mathrm{TR}$), then $\frac{1}{T_1^*}$ can be expressed by

$$\frac{1}{T_1^*} = \frac{1}{T_1} - \frac{\ln(\cos\theta)}{\mathrm{TR}}, \tag{2.11}$$

where θ is the applied flip angle (FA) and TR is the time interval between two successive data acuquisition. At the steady state, M_{ss} is a function of M_0, TR, T_1, θ, given by

$$M_{ss} = \frac{M_0(1 - e^{-\frac{\mathrm{TR}}{T_1}})}{1 - e^{-\frac{\mathrm{TR}}{T_1}} \cos\theta} \tag{2.12}$$

For DESPOT1, the steady state signal is similar to (2.12). Now we consider the transverse signal instead:

$$M_{ss} = \frac{M_0(1 - E_1)\sin\theta}{1 - E_1 \cos\theta}, \tag{2.13}$$

where $E_1 = e^{-\frac{TR}{T_1}}$.

Keeping the TR constant and increasing the flip angle θ generate a curve characterized by T_1. All the data can be expressed in the linear form $Y = mX + b$

$$\frac{M_{ss}}{\sin\theta} = E_1 \frac{M_{ss}}{\tan\theta} + M_0(1 - E_1), \tag{2.14}$$

from which the slope m, and the y-intercept b can be estimated by linear regression. Therefore, T_1 and M_0 can be extracted by

$$T_1 = -\frac{TR}{ln(m)}, \tag{2.15}$$

$$M_0 = \frac{b}{1 - m}. \tag{2.16}$$

In the next section, the DFA method is discussed in detail.

2.4 Characterizing the Accuracy and the Precision of the T1 Mapping with Dual Flip Angle Method Using the Monte Carlo Simulation

2.4.1 Introduction

Fast T1 mapping is a desired MRI technique for the longitudinal or cross-sectional characterization of lesion structures. Inversion recovery (IR) is considered as the

gold standard for T1 measurements. The associated long acquisition time hinders its practical use, especially for applications that require 3D high spatial resolution and broad volume coverage. The dual-flip-angle (DFA) method [16, 17] is proposed as a rapid alternative, as it is based on the steady state signal and hence long waiting time is not required. Although this method is known to be sensitive to flip angle errors [18] and incomplete spoiling of transverse magnetization [19], approaches can be taken to mitigate these issues. As the DFA technique is simple for both data acquisition and T1 estimation, knowledge of its accuracy and precision at different signal-to-noise ratio (SNR) levels is desired to allow it to be more widely accepted. In this section, DFA is compared with IR using numerical Monte Carlo simulations under influences of various noise levels.

2.4.2 Method

From the previous study on the DFA method [17], the transverse signal intensity S is a function of the longitudinal relaxation time T_1, the repetition time TR, the flip angle α, and the equilibrium longitudinal magnetization M_0:

$$S = \frac{M_0(1 - e^{-TR/T_1})sin(\alpha)}{1 - e^{-TR/T_1}cos(\alpha)}. \tag{2.17}$$

Instead of rearranging the formula into a linear form and calculating the fitted T_1 and M_0 by the y-intercept and slope, respectively, the approximate parameter method [20] is used in this simulation:

$$T_1 = \frac{2TR(S_1/\alpha_1 - S_2/\alpha_2)}{S_2\alpha_2 - S_1\alpha_1}, \tag{2.18}$$

$$M_0 = \frac{S_1S_2(\alpha_2/\alpha_1 - \alpha_1/\alpha_2)}{S_2\alpha_2 - S_1\alpha_1}. \tag{2.19}$$

Here, we assume that $M_0 = 100$, $TR = 25$ms, and two flip angles, 6° and 20° are used.

All numerical simulations are implemented in MATLAB 2020B. The Monte Carlo simulation is applied in order to consider noise in the calculations of T_1 and M_0. In

our implementation, 10,000 values of random noise are generated by the Gaussian distribution with a mean of 0 and a standard deviation of M_0 divided by the SNR, the SNR is increased from 80 to 800 with an increment of 80, and T_1 values are set from 600 ms to 2000 ms with an increment of 100 ms.

For IR, we use five inversion delays, 100 ms, 500 ms, 1200 ms, 2500 ms, 4000 ms and a post-acuquisition delay of 2000 ms. The M_0 value and the T_1 range are the same as in DFA, but the SNR is changed from 10 to 100 with an increment of 10. The Levenberg–Marquardt algorithm [21] is commonly used for solving non-linear least squares problems. Two important indicators, the accuracy and the precision, are calculated by the mean of relative error and by the coefficient of variation $\frac{\text{std}}{\text{mean}}$, respectively, where std is the standard deviation.

2.4.3 Results

Figure 2-4 shows the mean of relative errors and the coefficient of variation of two estimated parameters T_1 and M_0, using IR under different noise levels and T_1 values. The contour lines indicate the 5% and 10% values. Figure 2-5 shows the same indicators with the DFA method. For each fitting, a higher SNR guarantees higher accuracy and precision.

Under the same SNR range, the IR-derived values of T_1 and M_0 have a low precision than accuracy, whereas the DFA-derived values of T_1 and M_0 admit approximate equal precision and accuracy. For SNR higher than 50, the accuracy and the precision of T_1 and M_0 with IR are both under 5% for most T_1 values. Such an SNR threshold for DFA is 240. Since the image acquisition time for IR used in this simulation is about 64 times that of the DFA method with the same image resolution, IR needs to be compared to DFA in a way that the SNR in IR is multiplied by 8 for equivalent SNR efficiency in DFA. Thus the DFA method is more efficient to achieve the same accuracy and precision than the IR method.

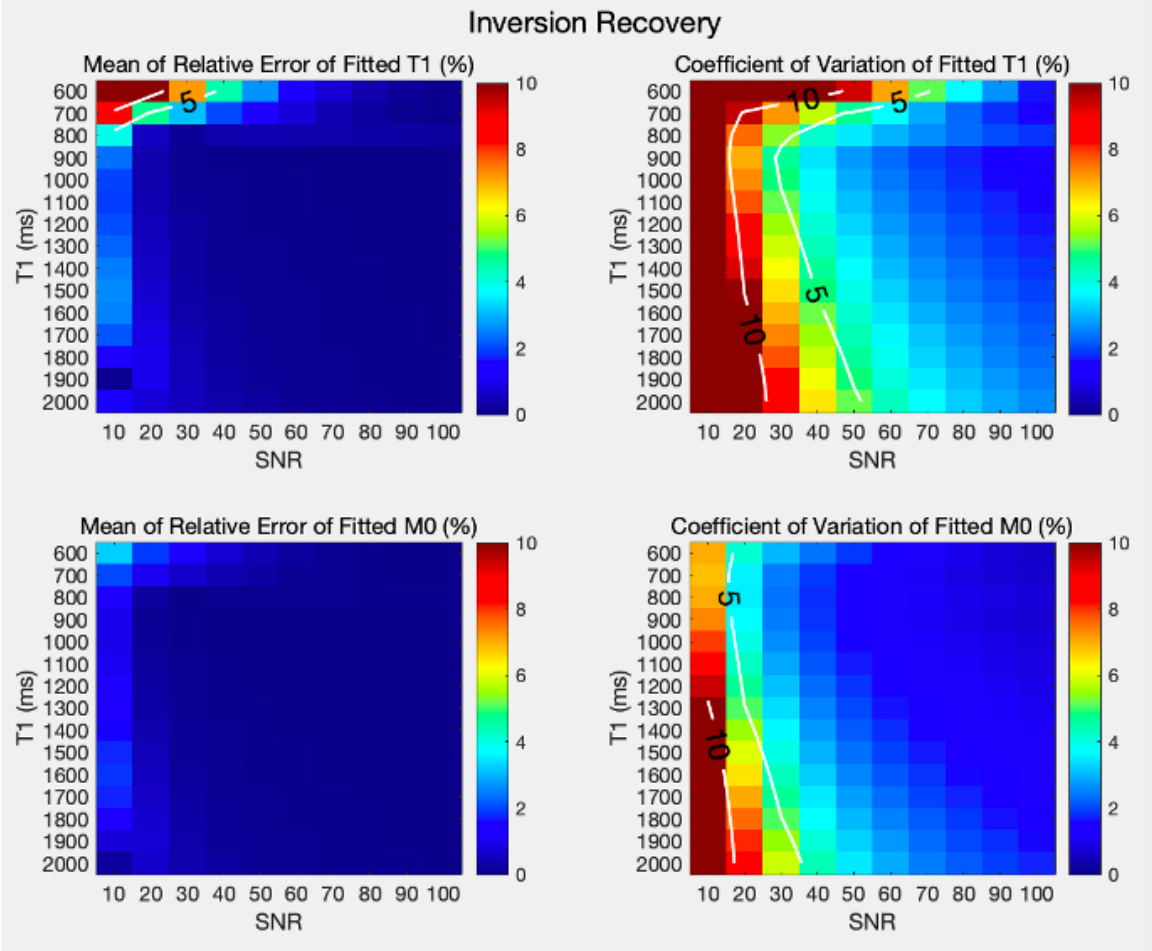

Figure 2-4. Simulation results of T1 and M0 estimation using IR. The mean of relative errors (indicating accuracy) and the coefficient of variation (indicating precision) of 10,000 repetitions as a function of the SNR using the Monte Carlo simulation.

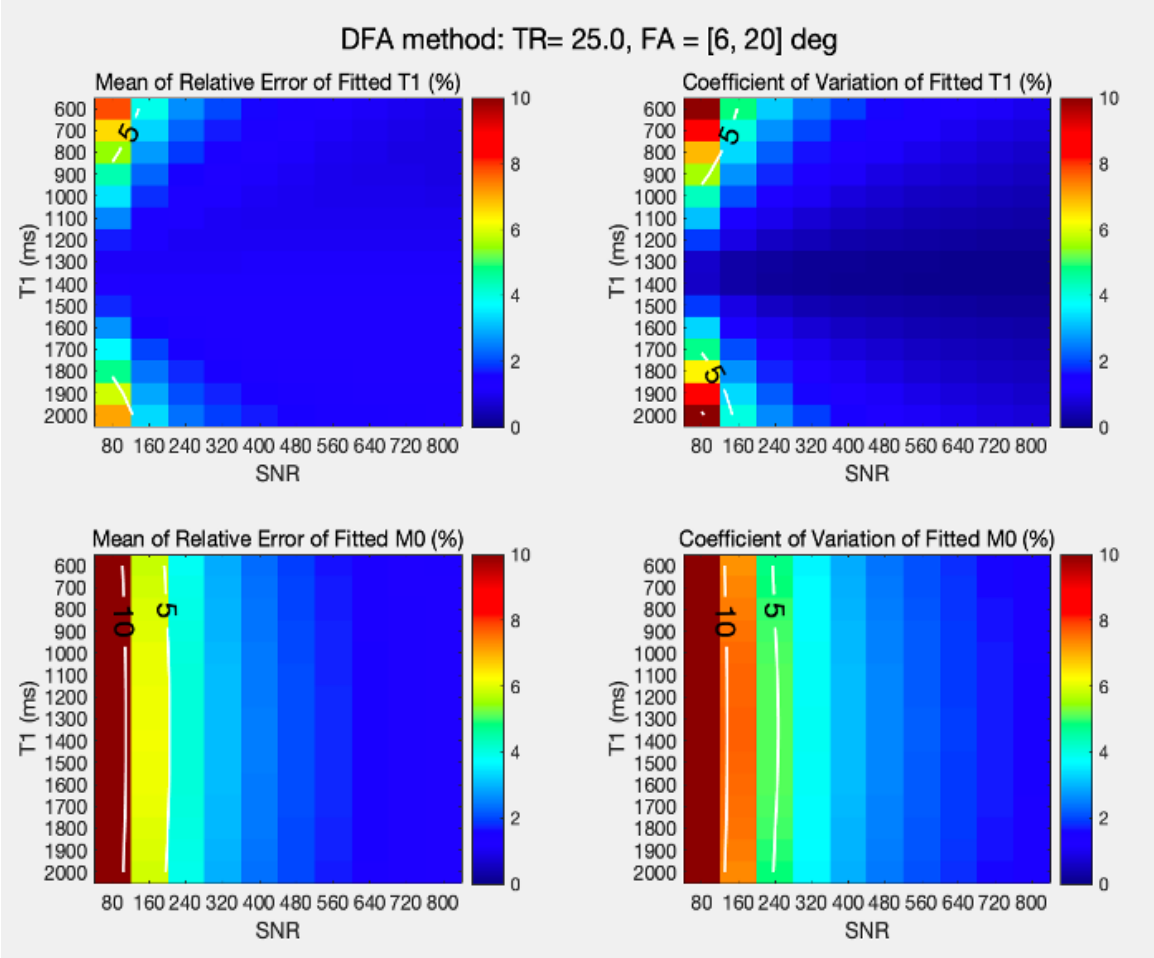

Figure 2-5. Simulation results of T1 and M0 estimation using the DFA method. Compared to Figure 2-4, it clearly shows that DFA has better performance and higher efficiency.

2.4.4 Conclusion

Our Monte Carlo simulation compared IR and DFA in a range of T1 values with different noise levels, and revealed that DFA has higher SNR efficiency for achieving the same level of accuracy and precision.

Chapter 3

Image Reconstruction

3.1 Introduction

Because MRI raw data are acquired in the k-space instead of the image space, image reconstruction is required for magnetic resonance imaging. Traditionally, MR images are reconstructed by simply applying inverse 2D or 3D fast Fourier transform (FFT). However, the situation is much more complicated in reality; for example, the scan time should be taken into consideration. As a result, several classic fast parallel imaging reconstruction algorithms are invented, including sensitivity encoding (SENSE) and generalized autocalibrating partial parallel acquisition (GRAPPA). This chapter is a tutorial that includes several reconstruction tasks to enhance the understanding of basic knowledge in this field. All the tasks are implemented in MATLAB 2020b. A set of 2D fully-sampled k-space raw data (whose dimension is 200×200) acquired with 8 MR coils is used in all the tasks.

3.2 Reconstruction Tasks

3.2.1 Reconstructing a 2D Image From Multi-Channel K-Space Data

Before we delve into the details, we first introduce the concepts of parallel imaging and sensitivity mapping. MR imaging uses multiple receiver coils to collect data. The

spatial information of the MR signal is localized by known placement and sensitivity of the receiver coils. In general, any number of coils, for example eight in our case, can be arranged around an object volume and then used in parallel to image the entire volume. Each coil will be very sensitive to the portion of the object which is near to the coil, and much less sensitive to the portion farther away. Each individual coil is capable of creating an MR image of the entire object. The formula is as follows,

$$\hat{I}(\vec{r}) = I_i(\vec{r})/S_i(\vec{r}), \tag{3.1}$$

where $\vec{r}$ denotes an arbitrary position in the image, $I_i(\vec{r})$ represents the image acquired from an individual coil i, $S_i(\vec{r})$ denotes the sensitivity map from the same coil, and $\hat{I}(\vec{r})$ is the estimated original image using this method.

The quality of the image acquired using a single coil is low. The portion farther from the coil has a lower SNR than the portion nearer the coil. Fortunately, these individual images can be combined using coil sensitivity maps to produce an improved image, given by

$$\begin{cases} S_1(\vec{r}) \times I(\vec{r}) &= I_1(\vec{r}), \\ S_2(\vec{r}) \times I(\vec{r}) &= I_2(\vec{r}), \\ & \cdots \\ S_8(\vec{r}) \times I(\vec{r}) &= I_8(\vec{r}). \end{cases} \tag{3.2}$$

The sensitivity maps provide the spatial information by looking at the relative intensity of each coil. Because the number of equations is greater than that of the variables (e,g, row number $m = 8$, column number $n = 1$ and $m > n$), the least squares fitting method is used to solve this problem. If we represent the system of equations in the matrix form $Ax = B$, then

$$A = \begin{bmatrix} S_1(\vec{r}) \\ S_2(\vec{r}) \\ \cdots \\ S_8(\vec{r}) \end{bmatrix}, \qquad x = I(\vec{r}), \qquad B = \begin{bmatrix} I_1(\vec{r}) \\ I_2(\vec{r}) \\ \cdots \\ I_8(\vec{r}) \end{bmatrix}.$$

The general solution to the least-square-fitting is given by multiplying both sides of the equation by the conjugate transpose of A, i.e. $A^H Ax = A^H b$. Hence,

$$x = (A^H A)^{-1} A^H b, \tag{3.3}$$

where $(A^H A)^{-1} A^H$ is called the *pseudo-inverse* of A. In our case,

$$\begin{aligned} A^H &= \begin{bmatrix} \overline{S_1(\vec{r})} & \overline{S_2(\vec{r})} & \cdots & \overline{S_8(\vec{r})} \end{bmatrix}, \\ A^H A &= \overline{S_1(\vec{r})} S_1(\vec{r}) + \overline{S_2(\vec{r})} S_2(\vec{r}) + \cdots + \overline{S_8(\vec{r})} S_8(\vec{r}), \\ (A^H A)^{-1} &= \frac{1}{\overline{S_1(\vec{r})} S_1(\vec{r}) + \overline{S_2(\vec{r})} S_2(\vec{r}) + \cdots + \overline{S_8(\vec{r})} S_8(\vec{r})}. \end{aligned}$$

Therefore,

$$\begin{aligned} x = (A^H A)^{-1} A^H b &= \frac{\overline{S_1(\vec{r})} I_1(\vec{r}) + \overline{S_2(\vec{r})} I_2(\vec{r}) + \cdots + \overline{S_8(\vec{r})} I_8(\vec{r})}{\overline{S_1(\vec{r})} S_1(\vec{r}) + \overline{S_2(\vec{r})} S_2(\vec{r}) + \cdots + \overline{S_8(\vec{r})} S_8(\vec{r})} \\ &= \frac{\overline{S_1(\vec{r})} I_1(\vec{r}) + \overline{S_2(\vec{r})} I_2(\vec{r}) + \cdots + \overline{S_8(\vec{r})} I_8(\vec{r})}{|S_1(\vec{r})|^2 + |S_2(\vec{r})|^2 + \cdots + |S_8(\vec{r})|^2}. \end{aligned} \tag{3.4}$$

Without loss of generality, we scale the values of all sensitivity maps simultaneously. Hence, we can easily normalize the denominator of Equation (3.4) to be 1. Therefore, the final equation is given by

$$I(\vec{r}) = \overline{S_1(\vec{r})} I_1(\vec{r}) + \overline{S_2(\vec{r})} I_2(\vec{r}) + \cdots + \overline{S_8(\vec{r})} I_8(\vec{r}). \tag{3.5}$$

Now we have combined all the images from all the receiver coils to reconstruct the original image by using sensitivity maps. Then we would like to reconstruct the image without knowing the sensitivity map. Rearranging the Equation (3.1), we get

$$S_i(\vec{r}) = \frac{I_i(\vec{r})}{I(\vec{r})}, \tag{3.6}$$

$$\overline{S_i(\vec{r})} = \overline{\left(\frac{I_i(\vec{r})}{I(\vec{r})} \right)} = \frac{\overline{I_i(\vec{r})}}{\overline{I(\vec{r})}}. \tag{3.7}$$

Substituting the (3.7) into (3.5), we get

$$I(\vec{r}) = \frac{\overline{I_1(\vec{r})}}{\overline{I(\vec{r})}} I_1(\vec{r}) + \frac{\overline{I_2(\vec{r})}}{\overline{I(\vec{r})}} I_2(\vec{r}) + \cdots + \frac{\overline{I_8(\vec{r})}}{\overline{I(\vec{r})}} I_8(\vec{r}). \tag{3.8}$$

Multiplying both sides of the equation by $\overline{I(\vec{r})}$, we get

$$\begin{aligned} I(\vec{r})\overline{I(\vec{r})} &= \overline{I_1(\vec{r})} I_1(\vec{r}) + \overline{I_2(\vec{r})} I_2(\vec{r}) + \cdots + \overline{I_8(\vec{r})} I_8(\vec{r}), \\ |I(\vec{r})|^2 &= |I_1(\vec{r})|^2 + |I_2(\vec{r})|^2 + \cdots + |I_8(\vec{r})|^2, \\ I(\vec{r}) &= \sqrt{|I_1(\vec{r})|^2 + \cdots + |I_8(\vec{r})|^2|}. \end{aligned} \tag{3.9}$$

Therefore, without knowing the sensitivity maps, we can calculate the original image by combining all the 8 coils.

In the first task, we reconstruct a 2D MR image from the multi-channel k-space data in order to be familiar with several MATLAB functions frequently used in images reconstruction, including `ifftshift`, `ifft2` and `fftshift`. First, the function `ifftshift` is applied to the first and second dimensions of the raw data. Secondly, we apply the `iff2` function to transform the k-space data to the image space. Then, we calculate the square root of the sum of the squares (SSOS) to combine 8 coils, given in (3.9).

Finally, the `fftshift` function is applied to shift the zero-frequency component to the center of the spectrum. The reason is that if we use MATLAB to perform the fast fourier transform, the "center" of the spectrum (representing the zero frequency component) is the first and the last elements in the array, and the "end" of the spectrum (representing the highest frequency components) is located at the center of the array. The final result is given in Figure 3-1. The images acquired from each MR coil are also provided in Figure 3-2.

The sensitivity maps, given in Figure 3-3, can also be calculated by combining Equations (3.6) and (3.9),

$$\hat{S}_i(\vec{r}) = \frac{I_i(\vec{r})}{\hat{I}(\vec{r})} = \frac{I_i(\vec{r})}{\sqrt{|\hat{I}_1(\vec{r})|^2 + |\hat{I}_2(\vec{r})|^2} + \cdots + |\hat{I}_8(\vec{r})|^2}. \tag{3.10}$$

The symbol $\hat{x}$ indicates that the variable x is estimated by calculation. We never know what the true original image looks like, since all images we get are reconstructed by numerical calculations. The spatial positions of the 8 coils are illustrated in Figure 3-4.

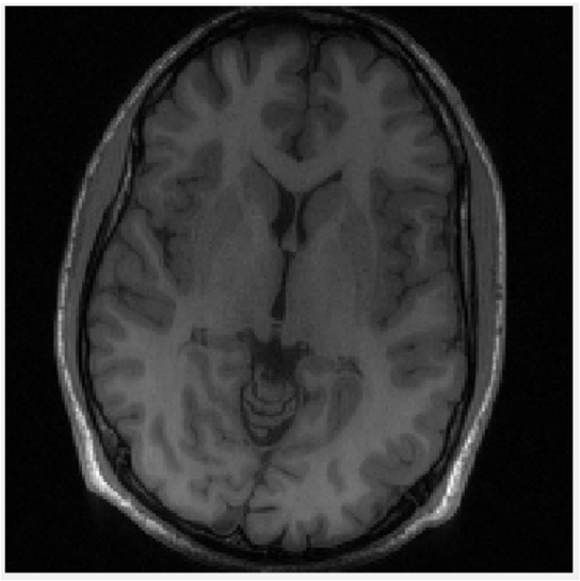

Figure 3-1. The reconstructed 2D MR image from the 8-channel k-space data.

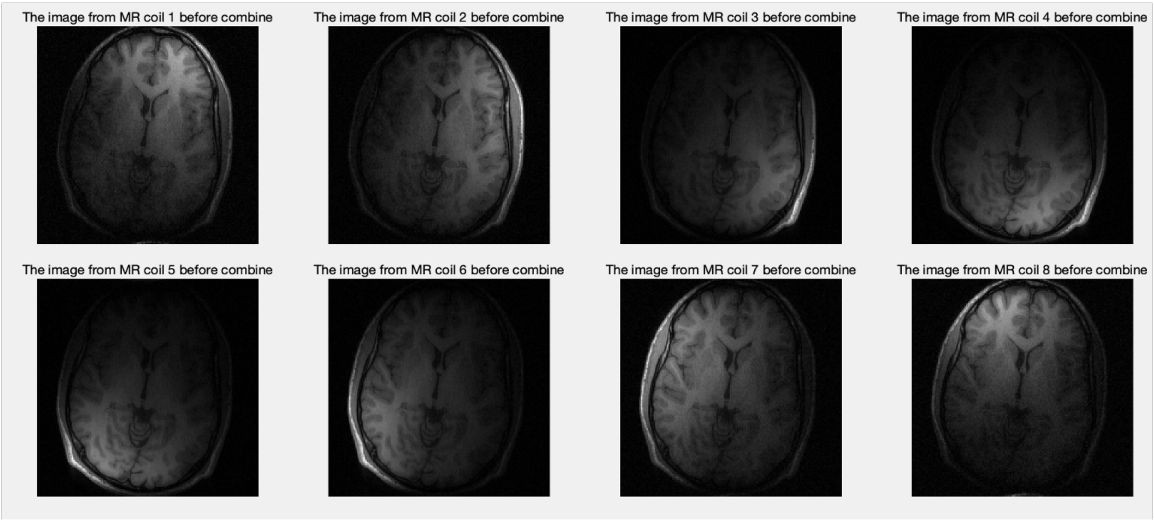

Figure 3-2. The images from each MR coil before the combination.

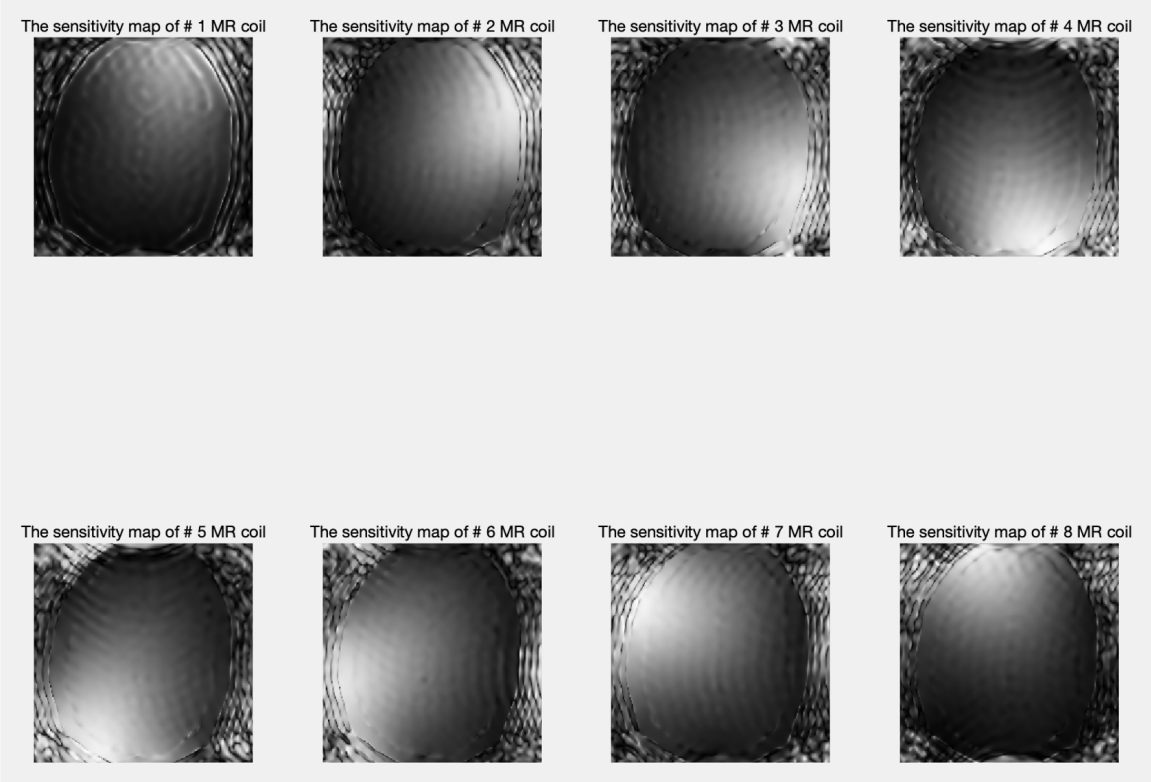

Figure 3-3. The calculated sensitivity map from each coil.

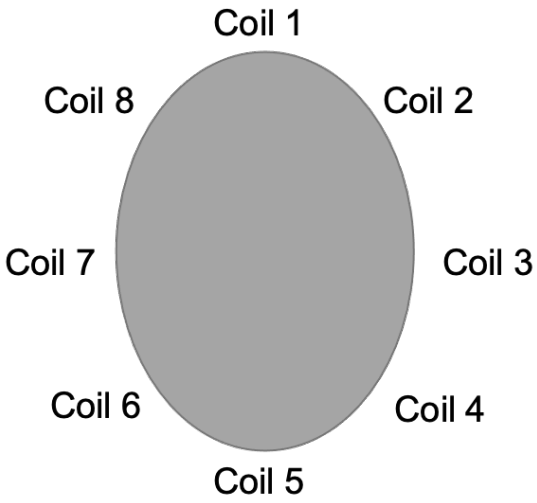

Figure 3-4. The spatial positions of the 8 coils.

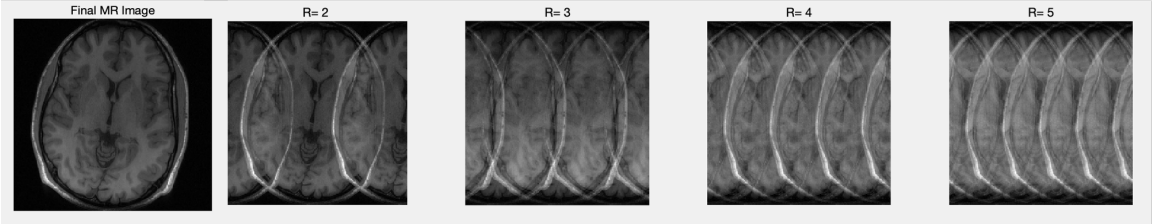

Figure 3-5. Equally spaced k-sapce undersampling images.

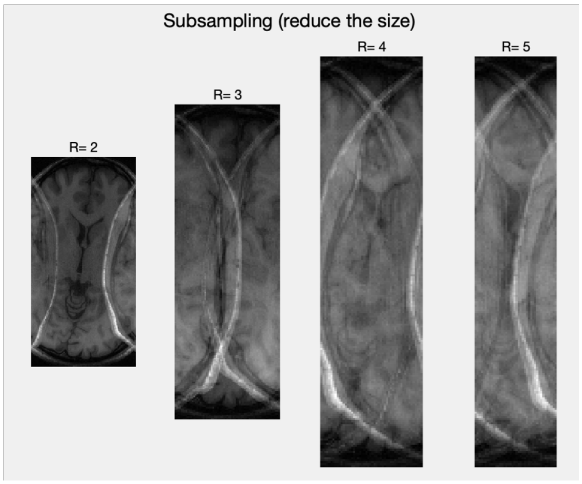

Figure 3-6. Subsampling images.

3.2.2 Equally Spaced K-Space Undersampling

Before learning about specific fast parallel imaging algorithms, it is essential to understand the concept of aliasing and basic undersampling techniques. In this task, five images are generated by equally sampling every other column (R=2), every third columns (R=3), every fourth columns (R=4) and every five columns (R=5) of the k-space's data. The result is given in Figure 3-5. Note that the signals are indistinguishable after undersampling.

In order to distinguish the undersampling and subsampling techniques, four images are generated by subsampling, given in Figure 3-6. The image size is reduced from 200×200 to 200×100 for R=2, to 200×67 for R=3, to 200×50 for R=4 and to 200×40 for R=5.

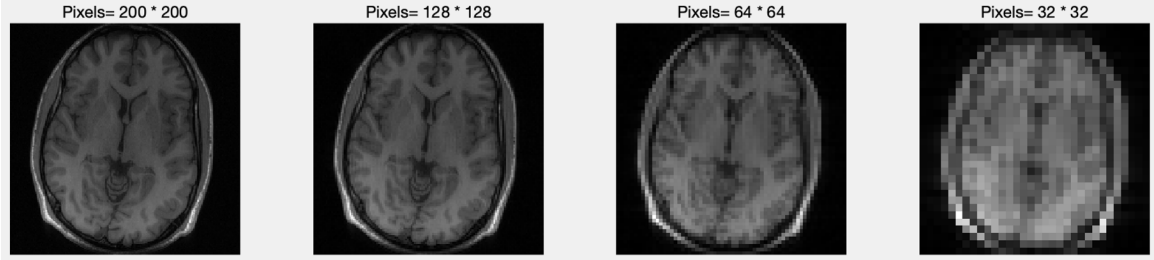

Figure 3-7. We truncate the image to 128×128 (resp. 64×64 and 32×32) pixels.

3.2.3 Image Resolution and K-Space FOV

3.2.3.1 Truncation and Interpolation

In Figure 3-7, we truncate the images from 200×200 pixels to $128 \times 128, 64 \times 64 and 32 \times 32$ pixels, respectively.

In Figure 3-8, instead of truncation, we keep the image size to be 200×200. We modify the image as follows. The central 128×128 (resp. 64×64 and 32×32) entries remain the same. However, the outer k-space entries are set to 0. If we want to resize the images from Figure 3-7 to 3-8, we need to apply image interpolation algorithms. *Interpolation* is a process to estimate unknown values according to surrounding known values. Image interpolation works in two directions, and tries to achieve a best approximation of a pixel's color and intensity based on the values of its neighboring pixels.

There are two categories of common interpolation algorithms: adaptive and non-adaptive. Adaptive methods incorporate the information of what is interpolated, e.g. sharp edges or smooth textures, whereas non-adaptive methods treat all pixels equally. Non-adaptive algorithms include nearest-neighbor, bilinear, bicubic, spline, sinc, Lanczos and others. The number of adjacent pixels used during interpolation varies from 0 to 256 (or more) depending on complexity. A larger number of adjacent pixels gives a higher accuracy but, on the other hand, requires much longer processing time.

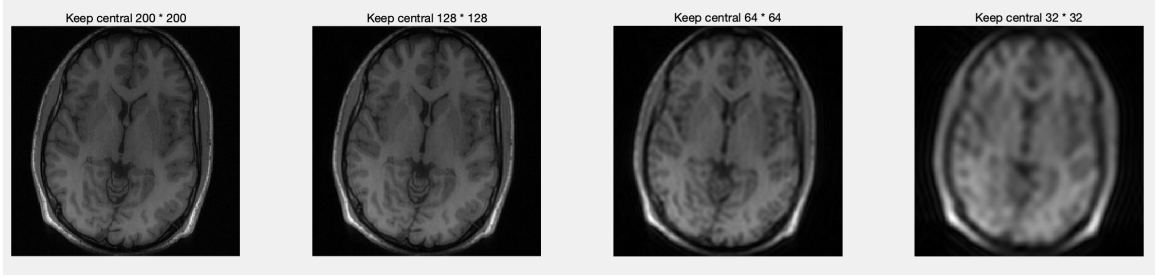

Figure 3-8. Instead of truncation, we keep the image resolution unchanged (200×200). The central 128×128 (resp. 64×64 and 32×32) entries remain the same but the outer k-space is set to 0.

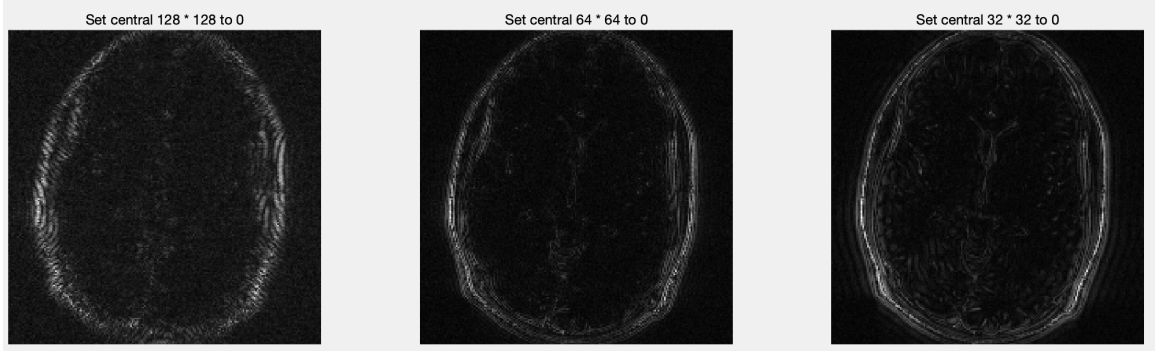

Figure 3-9. The central entries 128×128 (resp. 64×64 and 32×32) of k-space is set to 0 and the outer k-space remains unchanged.

Let us consider another situation: setting the central 128×128 (resp. 64×64 and 32×32) entries of the k-space to 0 and keeping the outer k-space unchanged. The results are given in Figure 3-9. Comparing this figure to Figure 3-8, we conclude that the low spatial frequency information is stored in the central k-space and the corresponding entries contain general intensity differences, i.e., the contrast information. Conversely, the outer k-space stores high spatial frequency information. The corresponding entries contain edge structures.

3.2.3.2 2D Hanning Window Design

In this section, we explain the design of the 2D Hanning filter to reduce ringing artifacts. In Figure 3-8, we keep the image size unchanged, the central 128×128 (resp. 64×64 and 32×32) entries remains the same and the outer k-space entries set to 0. If we zoom in on the 128×128 case, we can see some "echoes" near the edges, illustrated by red arrows in Figure 3-11. These are called *ringing artifacts*. In general, the cause of ringing artifacts is the ripple in the sinc function. In terms of the frequency domain, the signal passing a low-pass filter or a bandlimit filter is the main reason of ringing artifacts. (Recall that the sinc function is the impulse response for an ideal low-pass filter.) The overall results after applying the filter are given in Figure 3-10.

The *2D Hanning window*, also know as the *2D Hanning filter*, is one of the most commonly used filter. This filter has a bell shape curve. We take 32×32 case as an example to demonstrate the design in Figure 3-13. We choose the weight to be 4. The black square in this figure represents the matrix corresponding to the image, with dimension 200×200. Each entry in this matrix can be seen as a "point". Hence, the entries inside the blue circle correspond to the 32×32 submatrix that keeps the original values. The annulus between the red and blue circles corresponds to the transition part of this bell shape filter. In Figure 3-8, the value of an entry equals 0

outside the 32×32 submatrix and the original value inside the submatrix. Therefore, there is a large gap, from nonzero to zero on the boundary of the 32×32 matrix. The filter is designed to resolve this issue by smoothing the value differences in order to reduce ringing artifacts.

Let $x' = x * f(d)$, where x is the value of an entry inside the 200×200 matrix, x' is the corresponding value after applying the filter, f denotes the filter function, and d is the distance between the point and the central point $(100, 100)$. Every point inside the 200×200 matrix can be categorized into one of the following three cases:

1. If the point is inside the red circle (i.e., $d \leq 12$), then

$$f = 1.$$

2. If the point is outside the red circle but inside the blue circle (i.e., $12 < d < 16$), then

$$f = \cos^2(\theta),$$

   where θ is calculated by a linear function (by scaling the distance in the range [12, 16] to a radian in $[0, \pi/2]$).

3. If the point is outside the blue circle (i.e., $d \geq 16$), then

$$f = 0.$$

Entries are considered in the "transition stage" if they belong to the second case (i.e., used to fill the large gap between the nonzero values and the zeros). The result of the 128×128 case by adding the Hanning filter is shown in Figure 3-12. Comparing the areas indicated by the red arrows with those in Figure 3-11, it is obvious that the ringing artifacts have been remarkably reduced.

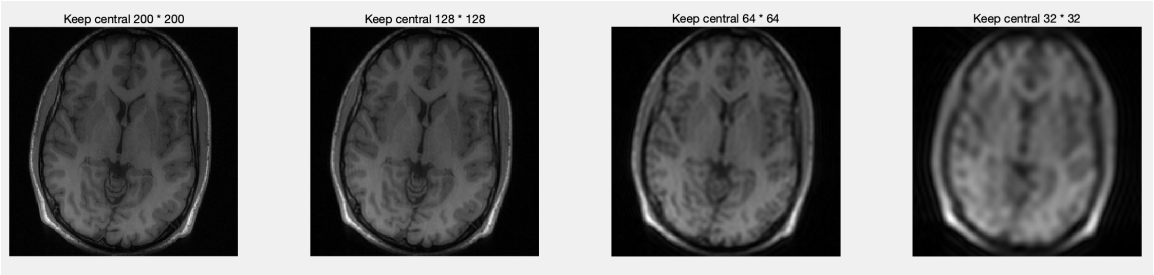

Figure 3-10. A 2D hanning window is added to the Figure 3-8.

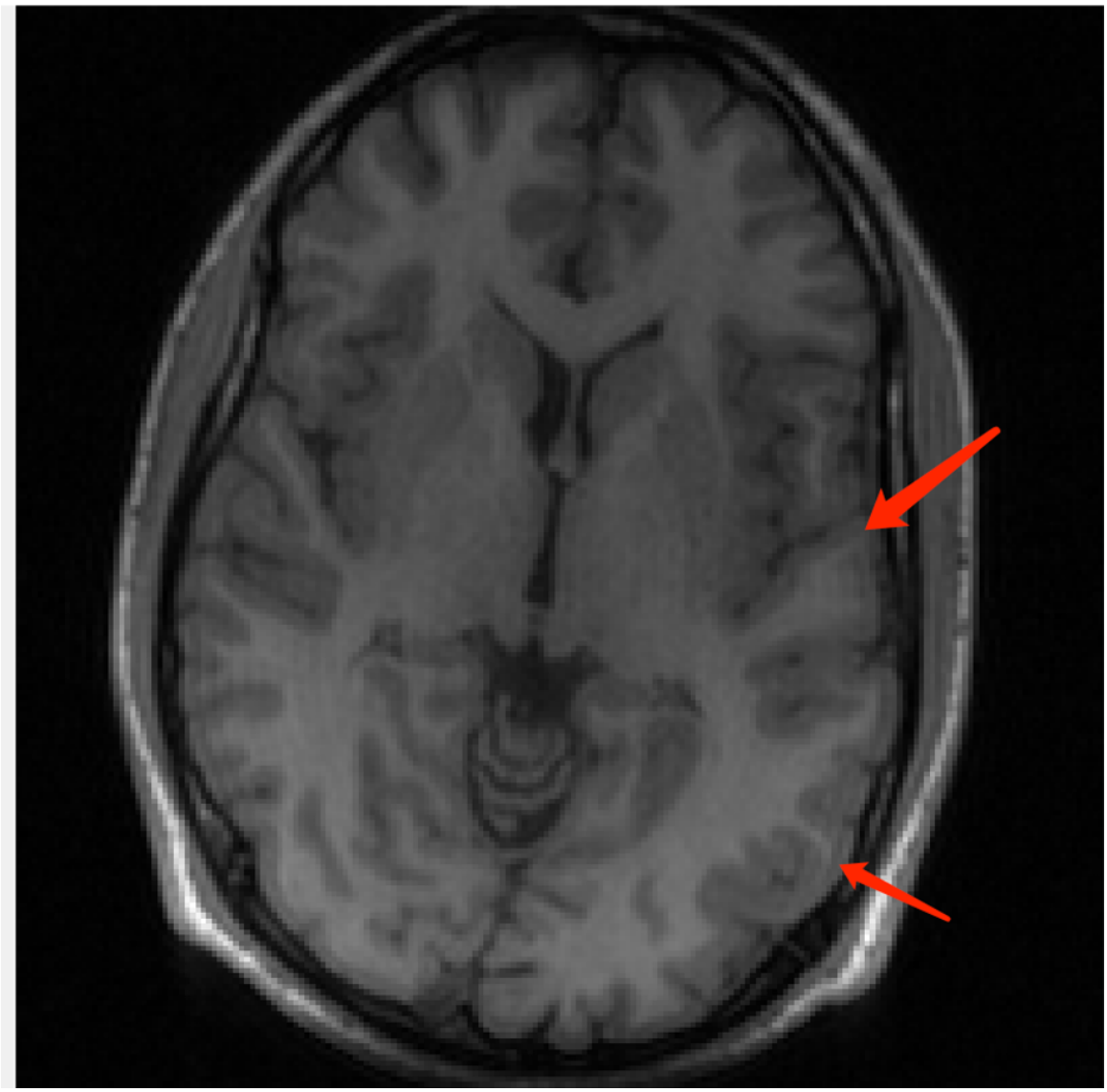

Figure 3-11. Zooming in on the 128×128 case from Figure 3-8 (i.e., without the 2D hanning filter).

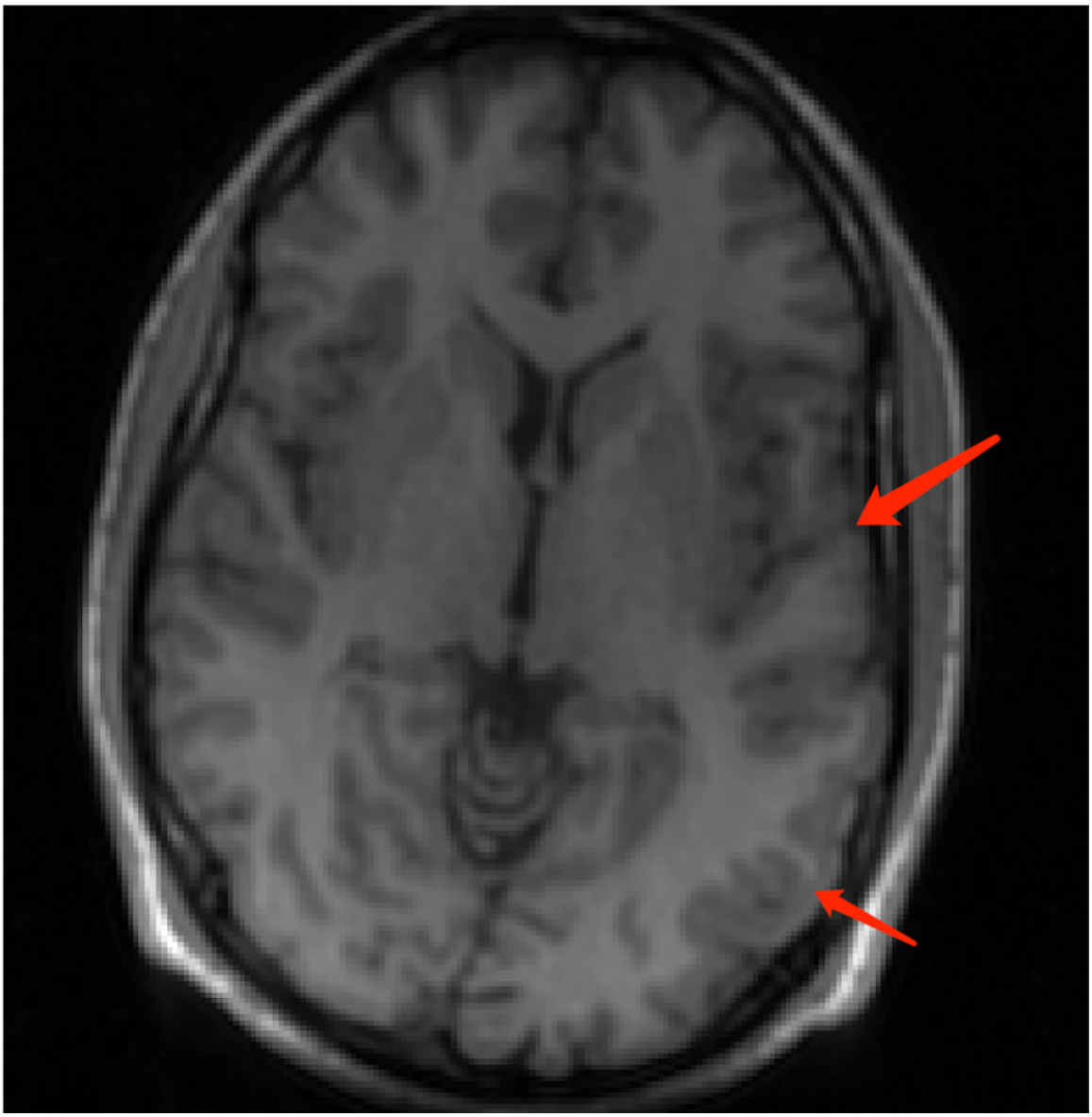

Figure 3-12. Zooming in on the 128×128 case from Figure 3-10 (i.e., with adding the 2D hamming filter).

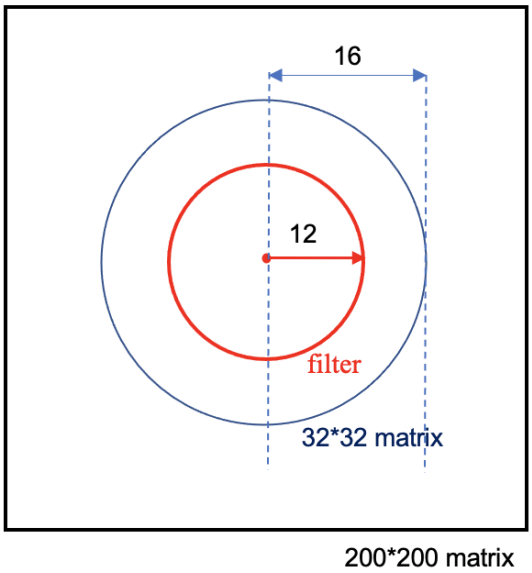

Figure 3-13. An illustration of the 2D hanning filter.

References

1. Hornak, J. P. *The Basic of NMR* 2014.
2. Bushberg, J. T. & Boone, J. M. *The essential physics of medical imaging* (Lippincott Williams & Wilkins, 2011).
3. Bitar, R. *et al.* MR pulse sequences: what every radiologist wants to know but is afraid to ask. *Radiographics* **26,** 513–537 (2006).
4. Dixon, R. L. & Ekstrand, K. E. The physics of proton NMR. *Medical physics* **9,** 807–818 (1982).
5. Mills, A. F., Sakai, O., Anderson, S. W. & Jara, H. Principles of quantitative MR imaging with illustrated review of applicable modular pulse diagrams. *RadioGraphics* **37,** 2083–2105 (2017).
6. De Deene, Y. *Review of quantitative MRI principles for gel dosimetry* in *Journal of Physics: Conference Series* **164** (2009), 012033.
7. Yokoo, T. *et al.* Estimation of hepatic proton-density fat fraction by using MR imaging at 3.0 T. *Radiology* **258,** 749–759 (2011).
8. Shah, B., Anderson, S. W., Scalera, J., Jara, H. & Soto, J. A. Quantitative MR imaging: physical principles and sequence design in abdominal imaging. *Radiographics* **31,** 867–880 (2011).
9. Goldman, M. Formal theory of spin–lattice relaxation. *Journal of Magnetic Resonance* **149,** 160–187 (2001).
10. Bloch, F. Nuclear induction. *Physical review* **70,** 460 (1946).
11. Bottomley, P. A., Foster, T. H., Argersinger, R. E. & Pfeifer, L. M. A review of normal tissue hydrogen NMR relaxation times and relaxation mechanisms from 1–100 MHz: dependence on tissue type, NMR frequency, temperature, species, excision, and age. *Medical physics* **11,** 425–448 (1984).
12. Chang, K. J. & Jara, H. Applications of quantitative T1, T2, and proton density to diagnosis. *Applied radiology* **34** (2005).
13. Chavhan, G. B., Babyn, P. S., Thomas, B., Shroff, M. M. & Haacke, E. M. Principles, techniques, and applications of T2*-based MR imaging and its special applications. *Radiographics* **29,** 1433–1449 (2009).
14. Chavhan, G. B., Babyn, P. S., Jankharia, B. G., Cheng, H.-L. M. & Shroff, M. M. Steady-state MR imaging sequences: physics, classification, and clinical applications. *Radiographics* **28,** 1147–1160 (2008).

15. Traficante, D. D. Relaxation. Can T2, be longer than T1? *Concepts in Magnetic Resonance* **3,** 171–177 (1991).

16. Deoni, S. C., Rutt, B. K. & Peters, T. M. Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* **49,** 515–526 (2003).

17. Helms, G., Dathe, H. & Dechent, P. Quantitative FLASH MRI at 3T using a rational approximation of the Ernst equation. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* **59,** 667–672 (2008).

18. Lee, Y., Callaghan, M. F. & Nagy, Z. Analysis of the precision of variable flip angle T1 mapping with emphasis on the noise propagated from RF transmit field maps. *Frontiers in neuroscience* **11,** 106 (2017).

19. Preibisch, C. & Deichmann, R. Influence of RF spoiling on the stability and accuracy of T1 mapping based on spoiled FLASH with varying flip angles. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* **61,** 125–135 (2009).

20. Blüml, S., Schad, L. R., Stepanow, B. & Lorenz, W. J. Spin-lattice relaxation time measurement by means of a TurboFLASH technique. *Magnetic resonance in medicine* **30,** 289–295 (1993).

21. Moré, J. J. in *Numerical analysis* 105–116 (Springer, 1978).

Appendix A

General Solution to a First Order Non-Homogeneous Linear Ordinary Differential Equation

Theorem A.1. *Given*

$$\frac{dy}{dx} = P(x)y + Q(x), \tag{A.1}$$

the general solution is

$$y = e^{\int P(x)dx}\left(\int Q(x)e^{-\int P(x)dx}dx + C\right).$$

Proof. Consider the special case where $\frac{dy}{dx} = P(x)y$ First, it is easy to get $\frac{dy}{y} = P(x)dx$. Hence,

$$\ln y = \int P(x)dx + C_0, \tag{A.2}$$

$$y = C_1 e^{\int P(x)dx}. \tag{A.3}$$

If we change C_1 from constant to a function related to x, i.e., $C(x)$, then

$$y = C(x)e^{\int P(x)dx}. \tag{A.4}$$

Differentiating Equation (A.3), we get

$$\frac{dy}{dx} = \frac{dC(x)}{dx}e^{\int P(x)dx} + C(x)P(x)e^{\int P(x)dx}. \tag{A.5}$$

Substituting (A.4) and (A.5) into (A.1) gives

$$\frac{dC(x)}{dx}e^{\int P(x)dx} + C(x)P(x)e^{\int P(x)dx} = P(x)C(x)e^{\int P(x)dx} + Q(x).$$

Then,

$$\frac{dC(x)}{dx} = Q(x)e^{-\int P(x)dx},$$

Therefore,

$$C(x) = \int Q(x)e^{-\int P(x)dx} + C \tag{A.6}$$

Substituting (A.6) into (A.4) gives

$$y = e^{\int P(x)dx}\left(\int Q(x)e^{-\int P(x)dx}dx + C\right).$$

□